Augmented Cognition: Allocation of Attention

Misha Pavel[1], Guoping Wang, Kehai Li
Oregon Health and Science University
OGI School of Science and Engineering
20000 NW Walker Road
Beaverton, OR 97006, USA
E-mail: pavel@ece.ogi.edu

ABSTRACT

We describe a novel, model-based approach to augmented cognition. We first discuss a subset of cognitive limitations that are likely to impair performance. We then note traditional approaches to the enhancement of cognitive abilities and their shortcomings, and compare the augmented approach to that of the traditional human-computer interface design. We discuss an example of our approach involving augmenting human ability to allocate attention in tasks such as search and monitoring. The model of attentional processes is then used to develop a framework for an optimal attention allocation process. Finally, we describe a pilot experiment in which observers' eye gaze is monitored and used to help them allocate attention.

1. Introduction

Tools that amplify human capabilities, and thus their performance, have been an important component underlying the progress of our society. The early, passive, tools simply extended human strength, while later machines enabled people to control energies and forces many times that of an individual. At the same time technology has been used to extend, in a limited way, human cognitive and communication capabilities. For example, written representation enhances memory and communication over distance. These aids, however, require extensive training. During the last 50-100 years, technology has been able to deliver increasingly more effective aids that extend our cognitive capabilities – such as calculators, and computer-based editors.

The early, passive tools were based on intuitive knowledge of the physics governing the situation. The major gains obtained when the engineers could use models of the situation and the task to optimize the tools. Similarly, the early – and many current – tools supporting cognitive tasks were based on intuitive knowledge of human cognitive limitations. The new, emerging approach to augmented cognition, however, is beginning to use models of cognitive mechanisms to overcome human limits, much like the engineers use physics.

The need to use underlying theoretical principles and models is even more critical in the development of tools for augmenting cognition than in amplifying human force because of the complexity of information spaces, and because the human abilities are likely to vary significantly across individuals, over time, and with the situational and environmental aspects of the task such as complexity of the task, workload, and fatigue.

In this paper we first note several areas where human cognitive abilities limit human performance, then we compare the new augmented cognition approach to that of traditional human-computer communication design, and finally we provide a specific example in augmenting human capabilities in allocation of attentional resources. The approach is based on our current understanding of attention allocation under uncertainty.

2. Background

2.1. Cognitive Limitations

Human information processing and decision making require a variety of diverse capabilities. Task analyses of cognitive tasks revealed requirements for cognitive

[1] This work was supported by the Augmented Cognition Program, DARPA Grant to OHSU number N66001-99-D-8503

0-7695-1874-5/03 $17.00 © 2002 IEEE

capabilities that underlie performance. These requirements include, but are not limited to:

- A large amount of precise working memory
- Ability to integrate multiple information sources
- Logical reasoning in complex and abstract problem spaces
- Incorporating uncertainties and values
- Enumerating possibilities
- Predicting consequences and discounting future values appropriately
- Appropriate distribution rather than focused attention – for anticipation of problems

Human abilities and their limitations have been explored by psychologists for many years. Although the thrusts of these experiments was to discover the underlying cognitive and physiological mechanisms, the resulting data and models can be used to describe quantitatively the cognitive limitations that affect human performance. One of the most obvious of these limitations is human memory, both short and long term. In addition to the limitation of human memory, however, are many other cognitive limitations, including logical reasoning abilities, difficulty with coordinate transformations, the ability to predict and anticipate states and actions, the fusion of information from abstract sources, and limitations of attention.

One of the most critical aspects of cognitive limitations involve human abilities to incorporate uncertainty in their judgment and decision-making. Although good-quality cognitive decisions are frequently made rather fast, decisions that require incorporation of uncertainties and values are very challenging. This is evidenced by numerous erroneous conclusions made by even sophisticated and statistically trained individuals. The numerous replicable errors in these situations have been called "cognitive illusions" because people make these mistakes even if they know about their propensity to make them. In that manner, these problems are similar to those where humans experience perceptual illusions that are not correctable by their cognitive knowledge of the ground truth.

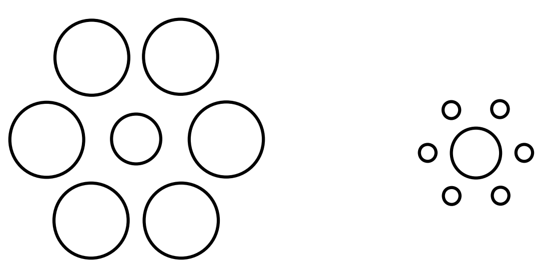

Figure 1. Example of a perceptual illusion. The two central circles are the same size. This illusion persists even if the viewer is aware of the physical reality.

An example of a perceptual illusion of size is shown in Figure 1. A viewer can convince himself by a simple measurement that the two central circles are the same size, but the illusion persists even after he or she performs the measurement. It is possible to design a perceptual aid that will "equalize" the perceived size of the central circles, but to do so requires fairly deep knowledge of the perceptual processes that lead to the illusion. Similarly, by starting with an understanding of the underlying cognitive limitation that leads to particular errors, we can design cognitive aids to reduce performance errors.

2.2. Traditional Solutions

As we noted above, finding ways to alleviate human cognitive limitations has been the subject of much design. Examples of such designs in various domains for alleviating the problems that arise due to the limitations in human memory include the following three approaches:

Memory Aids: Traditionally, human memory limitations have been addressed through storage and retrieval of information, using forms ranging from ancient symbolic representations to modern shorthand, checklists, mnemonic aids and variety of alerts. While effectively extending human memory capacity, memory aids themselves have limitations. First, information is often encoded in abstract forms, which requires substantial learning – a fairly difficult task for adults. Second, the sheer volume of information stored can make information retrieval difficult, both because of problems in indexing and in finding the written information when needed.

Training and Rote Learning: In order to improve performance in the context of stressful, and dynamic environments, human operators have been prepared by extensive training in more or less realistic contexts. The problem with training is that it takes a lot of time, and it is not a very efficient way of dealing with unpredictable situations.

Divide and Conquer: Critical tasks have been supported by designs where the tasks that could be performed by a single operator are simplified by dividing them among a number of individuals (examples: captain, first officer, navigator…) The shortcoming of the team-based approach to assuring cognitive capacity is the loss in efficiency and the problems that arise due to potential errors and difficulties in inter-personal communication.

In summary the shortcomings of these traditional approaches include costs, operational limitations, and limited gains. But the most critical shortcoming is the problem of generalization to new situations – and the potential inability to address unpredictable situations.

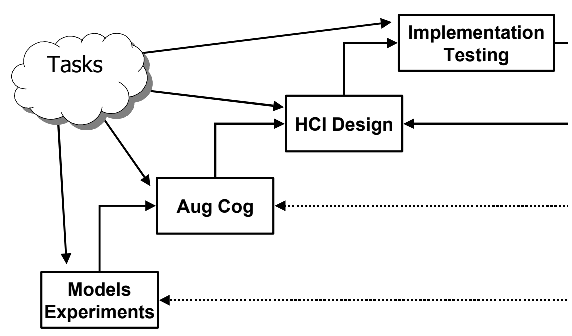

Figure 2. Human-Computer Interface (HCI) and Augmented cognition design process

3. Augmented cognition

3.1. Model-Based Approach

The proposed approach to alleviate the constraints due to cognitive limitations is to address the limitations directly. The starting point for the design based on augmented cognition is the characterization of the limitations to be alleviated. Ideally, these limitations would be specified using basic principles and represented in terms of information-processing models. The models could be then used to design the tools, much like physical principles and models are used to build mechanical tools. We expect, therefore, that for those cognitive abilities that have not yet been described by quantitative models, some basic research will be required. In the meanwhile, the limitation must be characterized by summarizing experimental data in terms of statistical or functional models.

Given the models, the designer may have an opportunity to develop devices that address the limitations directly by improving the information "impedance" match between the human cognitive system and the external information apparatus. For example, transformation of information may result in higher bandwidth of human information processing. We will discuss several examples that appear particularly suitable, but prior to that we would like to note the relationship between the augmented cognition and the traditional human-computer interface (HCI) design.

3.2. Comparison with HCI Design

One may argue that the traditional design of human-computer interfaces implicitly surmounts many cognitive limitations, and thus is already augmenting cognition. For example, the screen of a computer can be thought of as an extension of human working memory, and text editors may be viewed as amplifiers of human long-term memory. It is therefore appropriate to question the novelty of the augmented cognition approach.

The major difference involves the starting point and the objective of the design. The current user interfaces are typically designed to aid users in performing specific tasks. The design starts with an extensive task analysis followed by prototyping, and then by iterating extensive usability testing and prototype modification. In traditional HCI design, the developer attempts to develop interfaces that will match human capabilities. The developed interfaces, matched to the task and the user capabilities are, therefore, difficult to generalize. In addition, the gains in performance are then only marginal.

In contrast, the starting point for the augmented cognition approach is a specific cognitive limitation, ideally represented by a quantitative model. The designer then seeks a technological solution that alleviates the limitations. Such approach permits optimizing the solutions and is then generalizable to all tasks and systems that are limited by the same cognitive ability. The relationship is illustrated in Figure 2.

4. Example: Visual Attention Allocation

In this section we discuss a specific example where amplification of cognitive abilities is likely to lead to significant improvements in performance. One area where improvements in performance due to amplification of human cognitive capabilities would be most valuable involve situations where human errors lead to serious accidents, e.g., aviation.

Many errors and their consequences in a variety of mission-critical settings can be attributed to the operators' inappropriate allocation of attentional resources. This can be most clearly seen by analysis of accidents that are attributed to human error, and in particular, errors in aviation and medicine. In this paper, we present a simple approach to modeling attention allocation and suggest possible ways to aid human operators in more optimal allocation of processing resources.

4.1. Model of Attention Allocation

The theoretical framework for modeling attention is adapted from the work of Sperling and Dosher [6,7]. Numerous studies of attention allocation have demonstrated that, when a human operator is confronted with multiple tasks that require similar processing resources, there is a tradeoff in performance on these tasks as a function of the, typically voluntary, allocation of the resources. One way to characterize these tradeoffs between two tasks performed simultaneously is the so-called attention operating characteristics (AOC) depicted by the graph in Figure 3. In this graphs the points R_1 and R_2 represent the best possible performance on each task individually. The solid curve on this graph represents the

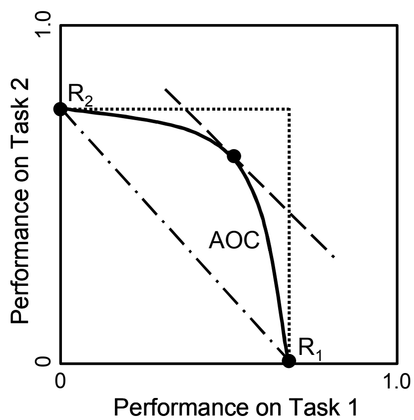

Figure 3. Attention operating characteristics. See the text for an explanation.

tradeoff between the performance on the two tasks and the particular shape of the curve reflects the degree to which the two tasks share resources.

In particular, if the tasks were independent of each other, and required separate resources, the curve would form a rectangle shown by the dotted line; i.e., the performance change in one task would not affect the performance on the other. In contrast, if the operator could perform only one of the tasks at a time, the AOC curve would represent the proportion of times that each task was attended and the resulting shape would be a straight line connecting the point of best performance on the individual tasks, i.e. R_1 and R_2.

In addition to the AOC curve, the interaction among the performances on multiple tasks can be characterized by the correlation or the conditional probability involving the separate tasks. The underlying idea is that if the performance on one task interferes with the other, this interaction would be reflected by a negative correlation among the performances. In those cases where the performance is all-or-none, the same information is captured by the contingency tables. In those situations, it is possible to test the degree of independence among the tasks by a statistical analysis of the contingency tables. The degree to which two tasks are negatively correlated indicates the resources that are switched among the tasks as opposed to shared.

The typical model of attention allocation in visual tasks that have been examined empirically is a combination of the switching and sharing resources. There are, however, tasks where switching is the dominant strategy. Examples of this situation include those tasks that require direct eye gaze. More specifically, in a variety of visual perceptual tasks, the high-resolution fovea can be allocated to only one location at a time. The tasks that require such high resolution, therefore, are limited by the ability of the operator to direct their gaze to the most appropriate location. A specific example of such task is visual search, or monitoring visual information in spatially disparate multiple locations. The remainder of this paper will be focused on the augmented cognition approach applied to visual search experiments.

4.2. Optimal Resource Allocation

In this section we use a simple "switching" resource allocation model of attention to illustrate the augmented cognition approach. We use visual search as the task to illustrate our approach. The theoretical framework is resource allocation, and the starting point is based on the AOC shown in Figure 3. The performance is characterized in terms of expected utilities: the probabilities $\{P_1, P_2\}$ of correct responses and the utilities associated with each correct response, $\{u_1, u_2\}$, assuming that incorrect responses have zero utility. The AOC can be then described as a function $P_2(P_1)$, and the expected utility over a number of trials is

$$E\{U\} = u_1 P_1 + u_2 P_2(P_1). \quad (1)$$

To maximize the expected utility we differentiate the expected utility with respect to P_1 and set the derivative to zero. The resulting optimal operating point (P_1^*, P_2^*) then satisfies

$$\frac{dP_2^*}{dP_1^*} = -\frac{u_1}{u_2}, \quad (2)$$

which means that the operating point is selected on the AOC curve at a point where the tangent of the AOC curve is the negative of the utility ratios as shown in Figure 3. Thus the optimal resource allocation depends not only on the performance tradeoffs characterized by the AOC curve, but also on the utilities associated with each response.

A realistic application of the optimal allocation approach requires that it be generalized to more than two tasks. A natural way to generalize the optimization approach described above is based on the theory of optimal search [3,4,5,8]. Let's assume that an operator is searching or monitoring *N* locations for an appearance of a target, and the target's prior probability at each location is given by the target prior probability P_i. If the operator allocates a_i of his attentional resources to location *i*, the probability of detecting a target at that location is $p(a_i)$. Let us further denote the utility of finding a target at the *i*-th location with resources by u_i. The operator's total cost, *C* is constrained by the total amount of available resources *A* so that

$$C = \sum_{i=1}^{N} a_i \le A. \qquad (3)$$

To compute the optimal resource allocation, we must maximize the expected utility associated with the detection of a target,

$$E\{U\} = \sum_{\forall a} P_i p(a_i) u_i, \qquad (4)$$

subject to the constraint in Equation (3). This optimization can be solved using the method of Lagrangian multipliers. The interesting and rather simple

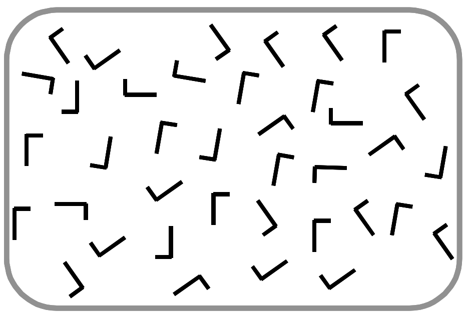

Figure 4. Example of a search paradigm. The target is the letter "L".

result is that the optimal resource allocation to each location is proportional to the rate of return from that location. These results are very similar to those for allocation of attention expressed by Equation (2).

4.3. Experimental Results

To examine the applicability of this approach to an experimental situation, we consider a particular task, in this case, searching for the letter "L" among its mirror images, illustrated in Figure 4. The stimuli were displayed on a computer-controlled CRT and subject's task was to indicate the location of the target. This is a perceptually difficult task – as opposed to the easier, pre-attentive tasks [9]. To find the target, the observer must examine a number of locations because he or she can make the detection decision only by looking in close proximity of the target. The difficulty of the task, therefore, depends on the number of the inverted "L" stimuli – the distracters.

We performed a pilot experiment with the stimuli similar to those shown in Figure 4. On each trial, the stimulus consisted of a single target and a number of distracters – the mirror images of the target. Each stimulus – target and distracters – were presented at random orientations. The stimulus remained visible until the observer found the target.

In this experiment we used the duration of the participant's gaze at a given location as the measure of the amount of resource allocated to that location. We assumed an exponential utility – detection – function

$$u_i(t_i) = 1 - e^{-\alpha t_i}$$

with the time constant of approximately 300 milliseconds, that was determined from the results of preliminary detection experiments. Given these assumptions, it was possible to estimate the instantaneous expected rate of return at each location in real time. We used these estimates to adjust the contrast of each stimulus item. In particular the contrast of each stimulus item was proportional the expected rate of return.

$$c_i \propto \frac{p_i u_i'(t_i)}{\sum_{j=1} p_j u_j'(t_j)}$$

Therefore, in the augmented cognition condition, the stimuli that were fixated longer would have lower contrast than the rest of the items. In that way, we were aiding the participants to allocate their attention to the stimuli with

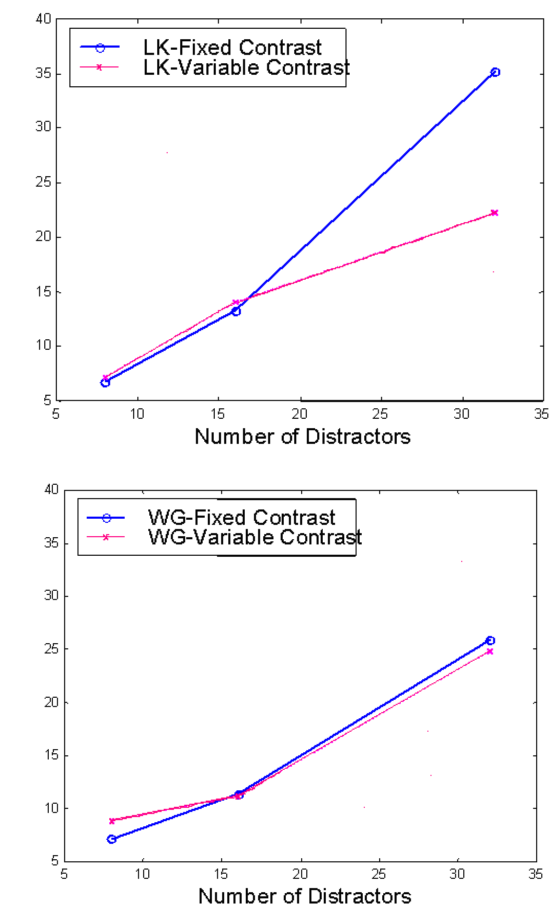

Figure 5. Results of the search experiment for two observers. The ordinate represents the response time required to find the target in seconds.

the highest instantaneous likelihood of target detection.

In order to assess the time allocated to each location, we monitored the eye gaze direction during each trial. The eye tracking was performed at a 60 Hz sample rate using an ASL 504 eye tracker, which estimates the gaze location by measuring the difference between the image of a pupil and a corneal reflection.

There were two conditions in this experiment. In the control condition the stimulus contrast was constant during the duration of each trial. In the augmented cognition condition, the contrast of the stimulus fixated by the observer was decreased. Thus, the observer was given perceptual information about where he had already allocated his resources and where he would need to look again. The results of this experiment for two subjects are shown in Figure 5. One of the observers showed considerable improvement while the other did not.

A subsequent analysis of the reason for the difference resulted in an important finding. The second observer's eye was much more difficult to track and the tracker often reported a location that was offset from the actual direction of gaze. In those instances, the reduction in contrast of the stimuli that had not been fixated, in fact, made the task more difficult rather than easier. We have upgraded our equipment to sample at a rate of 240 Hz, and are now gathering additional data.

5. Conclusions

We have proposed a new, model-based approach to augmenting the cognitive ability to allocate processing resources. Our goal was to develop techniques to alleviate limitations imposed by attention and processing capacity limitation in real time situation assessment and decision-making. Although our pilot experiment suggests there may be an advantage of the augmented approach in a specific situation, there is much to be done before we are ready to design routinely attention-enhancing tools to optimize human attention allocation and to incorporate uncertainty in real-time decision making. We have started a more comprehensive study using our current approach. In addition, we need to understand how to enable better control of attention allocation [1], and possibly include multimodal approaches [2].

6. Acknowledgements

We would like to acknowledge T.L. Hayes for her detailed review and editing assistance with this paper.

7. References

[1] Beck, J., Prazdny, K., and Rosenfeld, A. (1983). A Theory of Textural Segmentation. In *Human and Machine Vision*, Beck, J., Prazdny, K., and Rosenfeld, A., Eds., 1-39. Academic Press, New York.

[2] Bertelson, P. (1998). Starting from the ventriloquist: The perception of multimodal events. In: Sabourin, M., Craik, F., et al. *Advances in psychological science, Vol. 2: Biological and cognitive aspects, 419-439*. Erlbaum.

[3] Koopman, B. O., (1956) The theory of search: Target detection. *The Journal of the Operation Research Society of America*, 4, 503-531.

[4] Shaw, M.L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception & Performance*, 4, 586-598.

[5] Shaw, M.L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 201-211.

[6] Sperling, G. (1984). A unified theory of attention and signal detection. In R. Parasuraman and D. R. Davies (Hrsg.), *Varieties of attention* (S. 103-181)

[7] Sperling, G. and Dosher, B.A. (1986). Strategy and optimization in human information processing. In *Handbook of Human Perception and Performance: Volume 1*, K. Boff, L. Kaufman, and J. Thomas (eds). New York: Wiley.

[8] Stone, L. D., (1975). *Theory of optimal search.* Academic Press, New York.

[9] Triesman, A. (1985). Preattentive Processing in Vision. *Computer Vision, Graphics, and Image Processing 31*, 156-177.